Can Complexity Science Support the Engineering of Critical Network Infrastructures?

David L. Alderson
Department of Operations Research
Naval Postgraduate School
Monterey, CA 93943 USA
dlalders@nps.edu

John C. Doyle
Division of Engineering and Applied Science
California Institute of Technology
Pasadena, CA 91125 USA
doyle@cds.caltech.edu

***Abstract*— Considerable attention is now being devoted to the study of "complexity science" with the intent of discovering and applying universal laws of highly interconnected and evolved systems. This paper considers several issues related to the use of these theories in the context of critical infrastructures, particularly the Internet. Specifically, we revisit the notion of "organized complexity" and suggest that it is fundamental to our ability to understand, operate, and design next-generation infrastructure networks. We comment on the role of engineering in defining an architecture to support networked infrastructures and highlight recent advances in the theory of distributed control driven by network technologies.**

I. Introduction

Recent advances in computer networking technologies have greatly accelerated the extent to which system operators remotely monitor, manage, and control the physical world via the Internet and related communication systems. This is increasingly true for *critical infrastructure systems* (e.g., transportation, energy, telecommunication) at the local, regional, and even national levels. While this new ability has enabled great efficiencies in our economic, social, and civic lives, it has come at a price of new potential vulnerabilities to these systems. Public concern for the protection of critical network infrastructures against large-scale disruption has been prominent at the level of federal governments since the mid-1990s. At that time in the U.S., this concern was formalized through a Presidential Commission on Critical Infrastructure Protection (PCCIP), which was charged with assessing the extent to which national infrastructure systems were at risk to large-scale disruptions due to their interconnected nature as well as their growing dependence on the Internet as a "central nervous system" [1].

In the intervening decade, two trends have persisted in elevating the interdependence of the physical and cyber worlds as an important topic of study. First, our reliance on the Internet and other communication networks continues to grow rapidly, with a rich diversity of devices (e.g., mobile phones, PDAs, laptop computers) and applications (e.g., voice, email, text messaging, video) now integrated into daily use. Second, there is an increase in our collective sense of threat to these infrastructures from accidental failure (for example, the electric power outage in the Northeastern U.S. in August 2003), natural disasters (such as Hurricane Katrina) and deliberate attack (due to elevated terrorist activity worldwide). Such disruptions are grim reminders that even systems with historical reliability can have catastrophic vulnerabilities.

One crucially important, yet poorly understood, feature of national infrastructures, and complex systems in general, is that they are *robust yet fragile (RYF)*. In order to formalize this concept, we define *robustness* as follows.

Definition: *Robustness* is the invariance of a [*a property*] of a [*a system*] to [*a set of perturbations*].

Here, the use of square brackets emphasizes the notion that any formal definition of robustness at the systems level requires specification of the system, the property, and the set of perturbations. The concept of RYF underscores that a system can have a property that is robust for one particular set of perturbations, yet be *fragile* for [*a different property*] and/or [*a different perturbation*]. Indeed, a hallmark of both biology and advanced technologies is that they exhibit extremes of robustness and fragility, and highly evolved or well-designed systems effectively manage the resulting tradeoffs. Thus without specifying properties and perturbations, to say a system as whole is robust or fragile can only mean that the tradeoffs are handled well or poorly, respectively. A fragile system is one that has gratuitous fragility of most properties to most perturbations, while a robust one has targeted robustness of important properties to significant perturbations.

Robustness is a central issue in the study of critical infrastructures, and complex engineering systems in general, for several reasons. First, computer-based simulation is now powerful enough that it is relatively easy to create a demonstration of almost *anything*, provided that the circumstances are made sufficiently idealized. However, the real world is typically far from idealized, and thus a system must have robustness in order to close the gap between the demonstration and the real thing. A second closely related reason to study robustness is that most of the structural and behavioral complexity in many engineering systems is in order to provide this robustness in the first place.

Many system properties of interest can be viewed as a specific type of robustness. For example, *reliability* can be viewed as a type of robustness in which the set of perturbations include component failures. Similarly, *efficiency* can be interpreted as a type of robustness in which the set of perturbations include resource scarcity. Finally, *scalability*

Proceedings of the 2007 IEEE International Conference on Systems, Man and Cybernetics. Montreal, Canada, October 7-10, 2007.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
OCT 2001	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Can Complexity Science Support the Engineering of Critical Network Infrastructures?	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Postgraduate School Department of Operations Research Monterey, CA 93943**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
Proc. 2007 IEEE Internat. Conf. Systems, Man and Cybernetics. Montreal, Canada, October 7-10, 2007

14. ABSTRACT
Considerable attention is now being devoted to the study of complexity science with the intent of discovering and applying universal laws of highly interconnected and evolved systems. This paper considers several issues related to the use of these theories in the context of critical infrastructures, particularly the Internet. Specifically, we revisit the notion of organized complexity and suggest that it is fundamental to our ability to understand, operate, and design next-generation infrastructure networks.We comment on the role of engineering in defining an architecture to support networked infrastructures and highlight recent advances in the theory of distributed control driven by network technologies.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**8**	

can be viewed as a type of robustness in which the set of perturbations include changes to the size and complexity of the system as a whole. An example of RYF is that a system that is evolved toward high efficiency (i.e., robust performance with minimal system resources) might be very unreliable (i.e., it has fragile performance in the presence of even a single lost component). In many cases, a system design that supports a particular type of robustness directly leads to a different type of fragility.

A fundamental challenge facing the ongoing development of critical infrastructures is to organize system complexity in order to manage the tradeoffs between robustness and fragility. A prerequisite for this understanding is a clear notion of what it means for a system to be complex.

II. Notions of Complexity

There is no singular notion of complexity, and considerable confusion often arises due to different interpretations of this term. A comprehensive review of previous attempts to define complexity is beyond the scope of this short essay. However, we highlight several important concepts and use them as motivation for the discussion to follow.

A. Historical Perspective

Almost sixty years ago, Warren Weaver, then director of natural sciences of the Rockefeller Foundation in New York City, published an article in *American Scientist* titled "Science and Complexity" [2]. In this article, he contrasted three classes of problems facing science: *simple* problems, problems involving *disorganized complexity*, and problems involving *organized complexity*.

In Weaver's view, simple problems were those involving a small number of variables, such that they could be analyzed completely and with certainty. This was long before the discovery of chaos in deterministic, low-order systems, so a key element in Weaver's taxonomy is the number of variables under consideration. Nonetheless, Weaver attributes progress in the physical sciences—and their corresponding technological advances ranging from the telephone, to transportation via automobile and airplane, to hydroelectric power—during the seventeenth to nineteenth centuries to the successful application of the scientific method to simple problems.

However, he contrasts these achievements with a major shift in scientific thinking that was taking hold in the early twentieth century.

> *Rather than study problems which involved two variables or at most three or four, some imaginative minds went to the other extreme, and said: "Let us develop analytical methods which can deal with two billion variables." That is to say, the physical scientists, with the mathematicians often in the vanguard, developed powerful techniques of probability theory and of statistical mechanics to deal with what may be called problems of* disorganized complexity.

In describing disorganized complexity, Weaver's example is that of billiard balls, a system whose mechanics had received considerable attention at the time. Classical dynamics provide accurate descriptions of a small number of balls interacting on the table at once. As the number of balls increases to the order of dozens, then the computational requirements become cumbersome, and such problems were impractical to solve. But as the size of the table and the number of balls becomes very large, then the problem actually becomes easier, in the sense that the methods of statistical mechanics are applicable, and one can answer with precision certain questions related to average properties of the system.

However, it is important to note that Weaver used the term "disorganized" here to emphasize:

> *the methods of statistical mechanics are valid only when the balls are distributed, in their positions and motions, in a helter-skelter, that is to say a disorganized, way. For example, the statistical methods would not apply if someone were to arrange the balls in a row parallel to one side rail of the table, and then start them all moving in precisely parallel paths perpendicular to the row in which they stand. Then the balls would never collide with each other nor with two of the rails, and one would not have a situation of disorganized complexity.*

While Weaver acknowledges the prevalence of disorganized complexity in many important systems, he notes that there exists an intermediate class of problems between the extremes of simplicity and disorganized complexity.

> *The importance of this middle region, moreover, does not depend primarily on the fact that the number of variables is moderate... The really important characteristic of the problems of this middle region, which science has yet little explored or conquered, lies in the fact that these problems, as contrasted with disorganized situations with which statistics can cope, show the essential feature of* organization. *In fact, one can refer to this group of problems as those of* organized complexity.

Weaver's examples are primarily those of biological systems, yet the technology in the national infrastructures of today is rapidly approaching the complexity of simple biology. Should these systems be treated as organized or disorganized? As we will show below, this debate remains timely today.

B. Revisiting Simplicity

Before proceeding with a discussion of the *complex*, it is worth pausing first to reconsider what it means for a system to be *simple*. Due to the ever growing computational power of modern computers, it is now commonplace to solve systems with increasing numbers of variables, so Weaver's use of system size is no longer an appropriate characterization of simplicity. An alternate approach is to

say that a system is simple if it has "simple questions" and gives "simple answers." By *simple questions*, we mean that the questions of interest can be posed using models that are small and easy to describe and that they can be pursued using elegant experiments requiring minimal interpretation. By *simple answers*, we mean that theorems about models can be verified with short proofs and experiments have simple, reproducible outcomes that yield predictable results that are insensitive to small changes in model parameters. There are many classical examples of systems that have simple questions and give simple answers: the pendulum as a simple harmonic oscillator; simple RLC circuits; the interaction of two bodies via gravity; and simple Boolean logic circuits as implemented in much digital hardware. Many questions regarding their behavior are simple, as are the answers.

The triumph of reductionist science has been to reduce the *apparent complexity* of the world directly to an underlying simplicity. Physics has always epitomized this approach, and recently, molecular biology has successfully mimicked physics. Yet not all real systems can be treated as simple.

C. A Taxonomy for Complex Systems

Recognizing that there are widely divergent starting points to complexity from the fields of mathematics, biology, engineering, and physics, our aim here is a simple but universal taxonomy. We begin with a purely descriptive view, with the intent of contrasting several dramatically different perspectives of complexity as they have appeared in the literature. The objective is to bring an informed view of complexity to the topic of critical infrastructure systems.

In this section, we build upon the basic dichotomy first introduced by Weaver and present an enhanced view that contrasts much of the recent work in complexity science. Again, as a basis for comparison, we consider the extent to which a given system has "simple questions" and gives "simple answers."

"Chaocritical" Complexity. Some of the most profound insights of the last century relate to the idea that systems can have simple questions that do not have simple answers. A classic example from Newtonian physics asks to describe the behavior of interacting bodies via gravity. The motion of two interacting bodies is easily shown to yield periodic orbits (i.e., the behavior of the 2-body problem is simple), but the motion of even three interacting bodies can be chaotic and hard to predict long term (i.e., the behavior of the 3-body problem is not simple). Turing showed that easily described tasks like checking whether a computer program will halt (the classical halting problem) can be undecidable. Concepts related to undecidability, chaos, fractals, and critical phase transitions have dominated scientific thinking about complexity since the 1960s, with much of the emphasis on simple models that nevertheless yield complex behavior.

One response to this challenge to reductionist simplicity has been essentially to pursue what Weaver calls disorganized complexity. Examples include: *self-organized criticality (SOC)* [3]; *edge-of-chaos (EOC)* [4]; *scale-free networks (SFN)* [5] and much of "the new science of networks" (e.g., [6], [7]). While different in detail, models of SOC, EOC, and SFN share several features in common, reflecting their roots in statistical physics:

- They are based on *random ensembles* (respectively, of lattices, Boolean networks, graphs)
- that are *minimally tuned* (via a "friction" parameter, via correlation structure, via preferential attachment)
- to particular configurations that are "interesting" in the ensemble (a critical phase transition, a bifurcation point, or a power law node degree distribution).

The resulting condition (EOC, SOC, SFN) is fragile to everything but random rewiring or perturbations, to which it is perfectly robust, by construction. It is always fragile to targeted perturbations of any kind, and there is no attempt to create targeted robustness. A partial review of these properties, particularly the role of power law degree distributions, is available from [8].

We will refer to the complexity celebrated in this collective body of work as "chaocritical complexity" to emphasize the importance of simple, minimally tuned ensemble models giving rise to complex and fragile behaviors. While this kind of complexity is no doubt fascinating, it is entirely different from what is observed in many complex engineering systems, including critical infrastructures (the two meanings of "critical" here are unrelated, an unfortunate accident).

Organized Complexity. A different type of complexity occurs when relatively simple and robust answers arise in the context of complicated questions and/or models. Consider as an example the technology involved in the Boeing 777 aircraft. This vehicle is built from a parts list of over 1 million components, yet it has proven itself to be remarkably robust, in the sense that its design has enjoyed successful operation amidst varying weather and environmental conditions. This robustness is evidenced by the fact that there has not been a single fatality aboard this type of aircraft since its launch a decade ago. The triumph of the resulting organization in its design is that a system comprised of unreliable, uncertain, and changing components can work together in dynamic, uncertain, and hostile environments and with limited testing and experimentation to yield a predictable, reliable, adaptable, and evolvable system. This type of "organized complexity" requires carefully crafted interactions, either by design or evolution, as well as a completely different theory and technology from the study of either reductionist science or chaocritical complexity.

By design, organized complexity has more targeted robustness to specific perturbations in components and environment, yet it is almost always fragile to random perturbations in system configuration (unless specifically designed or evolved to be otherwise, for example, in ad hoc wireless networks). Systems exhibiting organized complexity are typically robust in terms of

- a system property that is carefully chosen to represent a particular system function; and
- a set of perturbations reflecting what is most common

and/or dangerous in the environment and component parts, or most effecting the property in question.

As always, uncertainty in systems exhibiting organized complexity is pervasive and may be modeled stochastically, but it still is very structured. In contrast, the chaocritical view largely ignores functional properties, structured uncertainty, and also an organized and structured response to make a system have robust performance.

Organized complexity of this type is a hallmark of highly evolved systems in both technology and biology. In fact, it is becoming increasingly clear that complexity in technological and biological networks is driven by control systems that manage the interaction among components, not the number or diversity of the parts themselves. In technology, this is evidenced by an explosion in complexity in computer networks, automobiles, airplanes, supply chains, package delivery, etc., where emphasis is increasingly placed on *where*, *when*, and *how* more than *who* and/or *what*. Similarly, in biology, a count of protein-coding genes is only weakly correlated with organized complexity of organism [9]. So organized complexity is not merely about robustness but about *the management of functional robustness and fragility*, with chaocritical complexity avoided as much as possible.

Irreducible Complexity. It is possible for systems to have both the complicated descriptions of organized complexity and the extreme fragility of chaocritical complexity, and little in the standard theory of complexity suggests how to avoid this potentially disastrous state. For example, in the context of biology, we might accumulate more complete parts lists but never "understand" how it all works. More relevant to next generation critical infrastructures, we might build increasingly complex and incomprehensible systems which will eventually fail completely yet cryptically. The potential for this type of "irreducible complexity" remains a serious threat to our ability to build a "next generation Internet" that fully integrates the physical and cyber world.

Note that irreducible complexity is not always undesirable, as in the case of cryptography, where the intention is to have something that it hard to simplify and fragile to perturbation. Yet for critical infrastructures, irreducible complexity is disastrous. As for biology, the extraordinary robustness and evolvability of the biosphere as a whole suggests that biology fundamentally is not irreducibly complex. Some properties of particular individuals that seem gratuitously fragile may be "frozen accidents" that have temporarily avoided selective elimination, and these may be intrinsically unpredictable. In general, the extreme and cryptic fragility of irreducibly complex systems, whether in biology or technology, suggest not intelligent design but the lack of it.

The aforementioned definitions yield a simple taxonomy for complexity in which we contrast, in one dimension, the descriptions and/or models (small vs. large) and, in another dimension, the system behaviors in response to perturbations in descriptions, components, or the environment (robust vs. fragile). Collectively, this leads us to the following view.

TABLE I
TWO DIMENSIONS OF COMPLEXITY

	small models	large models
robust behavior	simplicity	organized complexity
fragile behavior	chaocritical complexity	irreducible complexity

While the terminology for this taxonomy is tentative, we believe the categorization is not. When viewed in this unified framework, chaocritical complexity and organized complexity are opposites. Chaocritical complexity celebrates fragility, but organized complexity seeks to manage the inevitable tradeoff between robustness and fragility. While it seems that nearly all interesting complex systems are robust yet fragile, knowing how to identify and protect against system fragilities has been a source of confusion. This confusion is due, at least in part, to opposite notions of complexity and their corresponding models of system structure and behavior. We will review several examples in Section III, but first we demonstrate how the structure of many complex systems can be interpreted as a response to the constraints that are imposed on them.

D. A Constraint-Based View of System Complexity

A natural way to model any specific system, or class of systems, is to describe system structure in terms of the *constraints* that must be obeyed. For example, much of Newtonian mechanics can be described in terms of the relationship $F = MA$. We often like to write down these constraints as differential or algebraic equations, but they can be more general. Then, a key observation in the understanding of system complexity is that the features of most evolved systems are a consequence of specific constraints that are placed on their structure and/or behavior.

In the ideal case, one can find constraints that are simple to describe and whose consequences are simple to determine. But we know that even simple sets of constraints (e.g., 3 body gravitational dynamics, chaotic maps, etc.) can have consequences that are complex, and in particular so fragile as to be unpredictable (either mathematically or experimentally). In order to study system complexity further, it is useful to decompose constraints into 4 sources: system/environment, component, emergent, and protocol/architectural.

System-level constraints. Many complex systems, ranging from biological systems to critical infrastructures, have clear constraints on the needs of the system as a whole. These *system-level constraints* can include functional requirements (i.e., what the system needs to do), as well as the environmental and operating requirements (e.g., the conditions under which the system must need to be able to achieve this function).

Component-level constraints. At the same time, the components that comprise the system are typically constrained in terms of what they can do individually. That is, there are often physical, chemical, and/or informational *component-level constraints* to which the system must adhere.

The collective system can be viewed as a compromise between what is required of the system as a whole and what the individual components can do. This interaction of system-level constraints and component-level constraints often leads to the discovery of an additional type of constraint.

Emergent constraints. These are derived from the interaction between the system-level and component-level, and they dictate what is possible from a collection of individual components. These constraints are often expressed as hard limits, either absolute or in expectation, and are sometimes referred to as *laws*, the most well-known of which come from the fields of thermodynamics (Carnot), communications (Shannon), control (Bode), and computation (Turing, Gödel). Historically, knowledge about these hard limits has been fragmented, with each field making largely incompatible assumptions. However new unifications are encouraging. For example, recent efforts to integrate information theory and control theory [10] show considerable promise for providing new insight into what is possible in the realm of automated remote sensing and control.

Architecture. An "architecture" imposes additional constraints on the overall system, typically in the form of *protocols* or other rules for the configuration and/or interaction of system components. Although these additional constraints reduce the number of possible system solutions, a "good" system architecture constrains these solutions in a manner that focuses on the feasible solutions. System architecture will be the focus of Section IV-A.

We note that our consideration of question and answer complexity is not really new. In a sense, the study of computational complexity by computer scientists is really about the ratio of the complexity of answer/question in the limit when both go to infinity. The view presented here is more rudimentary.

III. Case Studies

Having introduced the contrast between chaocritical complexity and organized complexity, we review here several examples in which these opposing views result in dramatically different interpretations for the RYF nature of real systems.

A. Internet Topology at the Router-Level

The Internet serves as an important complex system for study from both a theoretic and practical perspective. One topic that has received considerable attention relates to the large-scale connectivity of the Internet at its different functional and organizational layers. After reports of power-laws in several types of network connectivity—including the router (physical) level, WWW (application) level, and Autonomous System (organizational) level—considerable attention has been placed on the development of models that replicate the empirically observed connectivity statistics. For example, one can use the tools of statistical mechanics to show that models based on *preferential attachment* during network growth, if tuned properly, yield power-laws consistent with empirical observation [11]. The resulting "scale free networks" (SFNs) suggest that the high connectivity nodes (in the tail of the power-law degree distribution) serve as central "hubs" that are crucial to the overall connectivity of the system and represent critical vulnerabilities if attacked [12]. In the spirit of "chaocritical complexity" these models celebrate the way in which simple rules give rise to complex behavior, with an emphasis on the fragility of the process that yields them and the overall fragility of the resulting structure.

Yet, in the case for models of the router-level Internet, SFNs have been shown to be a specious explanation for the appearance of power-laws in network connectivity and the RYF nature of the Internet in general [8], [13]. The basic reason is that technological and economic drivers of network design require that the network be *organized* to manage the tradeoff between system-level constraints on network performance and technological constraints on router throughput. Specifically, component-level constraints on router design (i.e., a conservation law in the number of packets that can be processed per unit time) give rise to an emergent constraint in the tradeoff between router connectivity and connection bandwidth (i.e., a router can have a few high bandwidth connections or many low bandwidth connections, but never an arbitrarily large number of high bandwidth connections) [14]. The resulting design is a sparse mesh of high-speed, low-connectivity routers in the network core, with high connectivity nodes only toward the network periphery, where they serve as access points that multiplex relatively low-speed end users, and not as central "hubs" in global connectivity. This structure not only provides high throughput performance, but it also results in considerable robustness to the loss of high connectivity nodes.

Thus, while models based on preferential attachment capture aggregate statistics and provide a superficial account of power-laws in the router-level network, they fall short of what is needed to understand this complex engineering system. Other aspects of the Internet, including traffic behavior and other types of network structure have been examined from these alternate perspectives, with similarly opposite outcomes [15].

B. Metabolic and Protein-Protein Interaction Networks

In microbiology, metabolic and protein-protein interaction networks serve as "critical infrastructure networks" at the cellular level. These systems are highly constrained at the systems level by highly unpredictable intracellular and extracellular environments as well as at the component level from physiochemical laws governing the individual parts (e.g., proteins and enzymes) and their reactions (i.e., conservation laws). Considerable attention has recently been directed at the structure of these network systems, with the discovery of power-law connectivity in both cellular metabolism and protein-protein interaction (PPI) networks cited as evidence that scale free networks are a fundamental structure in biology as well [16]. The claim is that when considering certain network representations of stoichiometry, the degree distribution for metabolite "nodes" is appropriately represented by a power law. An argument for SFNs parallel to

that for the router-level Internet has been made suggesting that metabolism is fragile to the loss of highly connected metabolites.

Yet, when viewed through the lens of organized complexity, in which domain-specific features of biochemistry are fundamental, one observes that an emphasis on power law connectivity in lieu of the domain-specific constraints at the system- and component-level leads to incorrect models of these biological networks, just as in the case of the Internet. When the statistics are done properly the available data strongly refutes the SFN models, and is more consistent with the ideas that biology is highly organized and optimizes tolerances and tradeoffs (HOT) [17] for efficiency, robustness, and evolvability [18], [19], [20].

C. Forest Fires

Real world fire-prone landscapes exhibit roughly power-law statistics in the size versus frequency of burned regions, and this feature has made forest fire modeling a popular topic within the complex systems community. Forest fires are perhaps the canonical application of SOC, which aims to explain the presence of power laws as the consequence of a system operating at a critical phase transition [21]. Once again, when the statistics are done properly the available data strongly refutes the SOC model, and are more consistent with HOT models that involve tradeoffs between robustness and fragility of forest yield to incident fires [22]. The contrast between SOC models and those based on engineering design has also been made in contexts other than forest fires [17].

There are two main HOT models of wildfires with different levels of resolution, but they both emphasize constraints and tradeoffs [22]. The component constraints are that the vegetation growth on long time scales and fire propagation on short time scales are highly constrained by biology and physics, as well as specifics of weather, climate, and topography. The system is largely robust to fires in the sense that most fires are small and the few large fires that dominate the statistics are rare, and vegetation has evolved to be specifically tuned to fire as a major disturbance. Ecosystems experts emphasize that ecosystems as a whole can evolve very rapidly due to "sorting," whereby competition and selection acts on a large pool of available organisms who are themselves undergoing slower lineage-based evolution. An important emergent property is that these constraints plus evolution lead naturally to state in which fires have a roughly power-law distribution, but very different from what is predicted by SOC models.

D. Internet Congestion Control

The Transmission Control Protocol (TCP) is the protocol responsible for managing network congestion (among other things) in the existing Internet architecture. TCP was developed in a somewhat ad hoc manner, based more on engineering intuition and experimentation than on any strict mathematical theory. Previous attempts to characterize TCP in terms of its aggregate statistics yielded a view that interpreted its behavior as a chaotic phenomenon [23]. In contrast, the view inspired by organized complexity is one in which application performance requirements must coexist with throughput constraints at the level of individual routes over the network. This latter view has recently led to the development of a rigorous mathematical framework showing that TCP (along with Active Queue Management, or AQM, at the routers) can be viewed as implementing a primal-dual optimization algorithm solving a global resource allocation problem [24], [25], [26].

The practical importance of this "discovery" is significant. Whereas previous efforts to validate network protocols focused on the simulation of realistic deployment scenarios, this approach facilitates theoretic proofs for networks that are arbitrarily complex (in terms of topology, number of routers and hosts, nonlinear behavior, and in the presence of delays) and yields short proofs of global stability in which the system equilibrium optimizes aggregate user utility (e.g., [27]). Collectively, these new theories of Internet congestion control and related networking technologies confirm engineering intuition and are yielding new approaches to the design and deployment of new protocols for high performance networking technologies, whether they be dramatic improvements to existing protocols [28] or the design of new cross-layer protocols [29], [30].

IV. The Role of Engineering and Design

While it is clear that different notions of system complexity can give rise to opposite interpretations of system structure and behavior, how does one determine what "matters" when it comes to real systems, such as critical infrastructure networks? Moreover, how does one move from the analysis of complex network systems to the deployment, configuration, and management of such systems? There has recently been considerable progress in our understanding of organized complexity and networks, derived from a revolution in the theory of distributed control driven by network technologies [10], [28], [29], [30], [31], [32]. This view of complexity focuses on organization, protocols, and architecture, and in this section we present an overview of several concepts that we believe will be fundamental to helping us understand, operate, and design next-generation network systems.

A. A Theory of Complex Network Architecture

What is architecture? In most usage, the term architecture focuses on the elements of structure and organization that are most universal, high-level, and/or persistent. It usually involves specification of protocols (rules of interaction) more than modules (which obey protocols). System architecture must facilitate system level functionality as well as robustness and evolvability to uncertainty and change in components, function, and environment. Architectures can be designed or evolve, but when possible should be planned. In the context of our constraint-based view of complexity, the role of architecture is to create new constraints (primarily in the form of protocols) that facilitate "good" solutions among competing component, systems-level, or emergent constraints.

To date, the study of systems architecture is more art than science. There are (at least) two approaches to the study of architecture. The first views architecture as *a set of design principles used to guide the construction of a system*, and as such focuses its research on the development of design principles. This is the approach taken in many systems engineering contexts [33]. The second approach is research on a particular system design and studies architecture as *the modularity, interfaces, functional decomposition, etc. that form the actual structure of the designed system*. For this latter view, the Internet again serves as a canonical example.

Much of the success of the Internet has been a result of adhering more or less faithfully over time to a set of fundamental network design principles adopted by the the early builders of the Internet [34] (e.g., layering, fate-sharing, end-to-end). In this sense, these principles constitute a modest "architecture" for the Internet. From today's perspective, this architecture is both brilliant in the choices that were made but shallow in our theoretical understanding of the full Internet protocol design problem, where engineering "design" has primarily taken the form of tinkering and/or intuition along with considerable experimentation. That is, the development of Internet technologies has followed from a largely empirical view, one in which *validation* of a design or protocol has been conducted via simulation or prototype. The success of this approach has resulted in a scenario in which we are better at "trial and error via deployment" than provable guarantees on performance, stability, etc. Moreover, it has perhaps given the false impression that the emergence of collective behavior is sufficient as a design outcome. However, as technological visions increasingly emphasize ubiquitous control, communications, and computing, with systems requiring a high degree of not only autonomy and adaptation, but also evolvability, scaliability, and verifiability, a more rigorous, coherent, and reasonably complete mathematical theory underpinning Internet technology is needed.

B. Progress for Networked Dynamical Systems

The study of networked dynamical systems is theoretically difficult for many reasons. One reason is that the interconnection of network components can be complex. Another reason is that, for systems of practical interest, the dynamics of the individual network components themselves can be complex. While much of the existing research has emphasized increased complexity along one dimension or the other, very little work to date has addressed the challenge of networked systems that have both complex interconnectivity and complex dynamics.

There exist many complex systems, both naturally occurring and man-made, that serve as existence proofs for what is possible in the design space of networked dynamical systems. Again, consider the Internet and biology as two canonical examples. Another is the flocking behavior on the part of birds and fish, long considered a classic example where collective system behavior cannot be predicted from the individuals, thus inspiring awe and wonder at, in the words of chaocritical complexity, the emergence of order from chaos [35]. However, recently there has been considerable progress in the use of a control-theoretic framework for understanding the organized complexity of flocking and synchronization in multi-agent systems, ranging from simple coupled oscillators to coordinated flight of unmanned aerial vehicles (UAVs) [36], [37], [38]. Until this work, and similar results in the Internet and biology, most attempts to "explain" the structure and behavior of these systems has utilized conjecture and simulation, with no theoretic "proof" to explain their success in function or performance. However, the last few years have witnessed considerable progress in the development of mathematical theory to show exactly why some complex networked systems work as well as they do and also to guide how to improve them. These results blend (from engineering) theories from optimization, control, information, and computational complexity, with diverse elements in areas of mathematics (e.g. operator theory and algebraic geometry) not traditionally thought of as applied (e.g., [39]).

Critical infrastructures are a particularly challenging domain for the application of systems engineering, in part because these systems have evolved in an ad hoc, piecemeal fashion, through a competitive landscape filled with merger, acquisition, and consolidation of individual infrastructures designed as stand-alone systems. While one clearly observes elements of design in the structure of these systems, it is sometimes necessary to reverse-engineer the drivers of this organization. In the meantime, the Internet remains a driving force in the integration of automated control in critical infrastructures.

V. Looking Ahead

A quick survey of the funded projects at the National Science Foundation (NSF) suggests that the Internet of tomorrow will support a diversity of devices, applications, and services [40]. Our dependence on the Internet is only going to increase, and this dependence will be amplified by a fundamental change in the way that we use the network. Current communications and computing is dominated primarily by human-to-human communication (e.g., email, chat, text messaging) or by human-to-machine communication (e.g., information retrieval, information storage, transaction processing). However, increased deployment of internet-enabled sensors and actuators that monitor and change the physical world means that the Internet is increasingly a platform for integrated control, communications, and computation. It is reasonable to believe that in the future a majority of Internet traffic will not directly involve a human and will be primarily between automated hardware and software. This transition will enable new capabilities and robustness, but will also expose new fragilities and vulnerabilities. The impact and implications of this type of change are not fully appreciated, and there exist few theoretical results that guarantee that efforts to implement a system of this magnitude will be successful.

Thus, the Internet has become critical in two ways. First, the Internet has become a type of public utility (like electricity or phone service) that underlies many important public

and private services. In this context, Internet disruptions have a "ripple effect" across the economy. Second, the Internet is increasingly a control system for monitoring and controlling our physical environment, and as a result, hijacking the Internet can be even more devastating than interrupting it.

What do we need in the next-generation critical infrastructures, including the Internet? Perhaps most fundamentally, we need to be able to manage the tradeoff between functional robustness and system fragility, so that we can appropriately balance the benefits of increased efficiency and convenience with the potential for large-scale disruption. We may not be able to eliminate the RYF nature of the Internet and other critical infrastructures, but we can minimize the potential risk of catastrophic failure.

Will complexity science support the engineering of next-generation infrastructure systems, including the Internet? If by "complexity science" one means an emphasis on the tools and techniques associated with the study of chaos-critical complexity, then we believe the answer is no. Our understanding of a systems-level architecture for complex networks is nascent, but it's clear that design of architectures is a topic of increasing importance. And while progress is being made in our understanding of organized complexity, considerable work remains. In particular, we cannot afford to wait to see what emerges from the ongoing integration of the Internet and other critical infrastructures, nor can we trust the validation of such systems to heuristic argument and simulation. Thus, there is a need for a new theory of network architecture, one in which the principles of organized complexity provide guarantees on the performance and reliability of the system as a whole. How to formalize such "proofs" remains an open area for research.

REFERENCES

[1] President's Commission on Critical Infrastructure Protection, "Critical Foundations," The White House, Tech. Rep., 1997.

[2] W. Weaver, "Science and complexity," *American Scientist*, vol. 36, pp. 536–544, 1948.

[3] P. Bak, *How nature works: the science of self-organized criticality*. Copernicus, 1996.

[4] S. Kauffman, *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, 1993.

[5] A.-L. Barabási, *Linked: The New Science of Networks*. Perseus Publishing, 2002.

[6] P. Ball, *Critical Mass: How One Thing Leads to Another*. Farrar, Straus and Giroux, 2004.

[7] M. Buchanan, *Nexus: Small Worlds and the Groundbreaking Theory of Networks*. Norton, W. W. & Company, Inc., 2003.

[8] L. Li, D. Alderson, J. Doyle, and W. Willinger, "Towards a theory of scale-free graphs: Definition, properties, and implications," *Internet Mathematics*, vol. 2, no. 4, pp. 431–523, 2005.

[9] R. J. Taft, M. Pheasant, and J. S. Mattick, "The relationship between non-protein-coding dna and eukaryotic complexity," *BIOESSAYS*, vol. 29, no. 3, pp. 288–299, March 2007.

[10] N. C. Martins, M. A. Dahleh, and J. C. Doyle, "Fundamental limitations of disturbance attenuation in the presence of side information," *IEEE Trans. on Auto. Control*, vol. 52, no. 1, pp. 56–66, January 2007.

[11] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. of Modern Physics*, vol. 74, pp. 47–97, 2002.

[12] R. Albert, H. Jeong, and A.-L. Barabási, "Attack and error tolerance of complex networks," *Nature*, vol. 406, pp. 378–382, 2000.

[13] J. C. Doyle, D. Alderson, L. Li, S. Low, M. Roughan, S. Shalunov, R. Tanaka, and W. Willinger, "The "robust yet fragile" nature of the internet," *Proc. Nat. Acad. of Sci. USA*, vol. 102, no. 41, pp. 14 497–14 502, 2005.

[14] D. Alderson, L. Li, W. Willinger, and J. Doyle, "Understanding internet topology: Principles, models, and validation," *IEEE/ACM Transactions on Networking*, vol. 13, no. 6, pp. 1205–1218, 2005.

[15] D. Alderson and W. Willinger, "A contrasting look at self-organization in the internet and next-generation communication networks," *IEEE Communications Magazine*, vol. 43, no. 7, pp. 94–100, July 2005.

[16] A.-L. Barabási and Z. Oltvai, "Network biology: understanding the cell's functional organization," *Nature Reviews Genetics*, vol. 5, pp. 101–114, 2004.

[17] J. M. Carlson and J. C. Doyle, "Complexity and robustness," *Proc. Nat. Acad. of Sci. USA*, vol. 99, pp. 2538–2545, 2002.

[18] R. T. T.-M. Yi and J. Doyle, "Some protein interaction data do not exhibit power law statics," *FEBS Letters*, vol. 579, no. 23, pp. 5140–5144, 2005.

[19] R. Tanaka, M. Csete, and J. Doyle, "Highly optimised global organization of metabolic networks," *IEE Proc. Systems Biology*, vol. 152, no. 4, pp. 179–184, 2005.

[20] R. Tanaka, "Scale-rich metabolic networks," *Physical Review Letters*, vol. 94, no. 168101, 2005.

[21] B. D. Malamud, G. Morein, and D. Turcotte, "Forest fires: An example of self-organized critical behavior," *Science*, vol. 281, pp. 1840–1842, 1998.

[22] M. A. Moritz, M. E. Morais, L. A. Summerell, J. M. Carlson, and J. C. Doyle, "Wildfires, complexity, and highly optimized tolerance," *Proc. Nat. Acad. of Sci. USA*, vol. 102, 2005.

[23] A. Veres and M. Boda, "The chaotic nature of TCP congestion control," in *Proc. IEEE INFOCOM 2000*, 2000.

[24] F. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.

[25] F. Kelly, "Mathematical modelling of the internet," in *Mathematics Unlimited - 2001 and Beyond*, B. Engquist and W. Schmid, Eds. Berlin: Springer-Verlag, 2001, pp. 685–702.

[26] S. H. Low and R. Srikant, "A mathematical framework for designing a low-loss, low-delay internet," *Networks and Spatial Economics, special issue on Crossovers between transportation planning and telecommunications*, vol. 4, pp. 75–101, Mar. 2004.

[27] A. Papachristodoulou, L. Li, and J. C. Doyle, "Methodological frameworks for large-scale network analysis and design," *ACM Comput. Commun. Rev.*, vol. 34, no. 3, pp. 7–20, 2004.

[28] D. X. Wei, C. Jin, S. H. Low, and S. Hegde, "FAST TCP: motivation, architecture, algorithms, performance," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 1246–1259, Dec. 2006.

[29] J. Wang, L. Li, S. H. Low, and J. C. Doyle., "Cross-layer optimization in TCP/IP networks." *IEEE/ACM Trans. on Networking*, vol. 13, pp. 582–595, 2005.

[30] L. Chen, S. H. Low, M. Chiang, and J. C. Doyle, "Cross-layer congestion control, routing and scheduling design in ad hoc wireless networks," in *Proc. IEEE INFOCOM 2006*, 2006.

[31] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition," *Proc. of the IEEE*, vol. 95, 2007.

[32] J. Doyle and M. Csete, "Rules of engagement," *Nature*, vol. 446, p. 860, 2007.

[33] E. Rechtin, *Systems Architecting: Creating and Building Complex Systems*. Prentice-Hall, 1991.

[34] D. D. Clark, "The design philosophy of the darpa internet protocols," *ACM Computer Communication Reviews*, vol. 18, no. 4, pp. 106–114, 1988, proc. ACM SIGCOMM 1988.

[35] S. Strogatz, *Sync: The Emerging Science of Spontaneous Order*. New York: Hyperion, 2003.

[36] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. on Auto. Control*, vol. 48, no. 6, pp. 988–1001, June 2003.

[37] A. Papachristodoulou and A. Jadbabaie, "Synchronization in oscillator networks: Switching topologies and presence of nonhomogeneous delays," in *Proc. ECC-CDC 2005*, Seville, Spain, 2005.

[38] A. Jadbabaie, N. Motee, and M. Barahona, "On the stability of the kuramoto model of coupled nonlinear oscillators," in *Proc. American Control Conference (ACC 2004)*, 2004.

[39] P. A. Parrilo, "Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization," Ph.D. dissertation, California Institute of Technology, Pasadena, CA, May 2000.

[40] Search www.nsf.gov for grants under the Computer & Information Science & Engineering Directorate.